

Novel application of a quantitative spatial comparison tool to species distribution data



Esther L. Jones^{a,b,*}, Luke Rendell^a, Enrico Pirotta^c, Jed A. Long^d

^a Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews KY16 8LB, United Kingdom

^b Centre for Research into Ecological and Environmental Modelling, The Observatory, Buchanan Gardens, University of St Andrews, St Andrews KY16 9LZ, United Kingdom

^c School of Mathematics, Washington State University, Vancouver 98686, WA, USA

^d Department of Geography and Sustainable Development, School of Geography and Geosciences, Irvine Building, University of St Andrews, St Andrews, Fife, Scotland KY16 9AL, United Kingdom

ARTICLE INFO

Article history:

Received 5 January 2016

Received in revised form 11 May 2016

Accepted 31 May 2016

Available online 15 June 2016

Keywords:

Edge effects

Map comparison

Moving window

Sperm whale

SSIM index

Uncertainty

ABSTRACT

Comparing geographically referenced maps has become an important aspect of spatial ecology (e.g. assessing change in distribution over time). Whilst humans are adept at recognising and extracting structure from maps (i.e. identifying spatial patterns), quantifying these structures can be difficult. Here, we show how the Structural Similarity (SSIM) index, a spatial comparison method adapted from techniques developed in computer science to determine the quality of image compression, can be used to extract additional information from spatial ecological data. We enhance the SSIM index to incorporate uncertainty from the underlying spatial models, and provide a software algorithm to correct for internal edge effects so that loss of spatial information from the map comparison is limited. The SSIM index uses a spatially-local window to calculate statistics based on local mean, variance, and covariance between the maps being compared. A number of statistics can be calculated using the SSIM index, ranging from a single summary statistic to quantify similarities between two maps, to maps of similarities in mean, variance, and covariance that can provide additional insight into underlying biological processes. We demonstrate the applicability of the SSIM approach using a case study of sperm whales in the Mediterranean Sea and identify areas where local-scale differences in space-use between groups and singleton whales occur. We show how novel insights into spatial structure can be extracted, which could not be obtained by visual inspection or cell-by-cell subtraction. As an approach, SSIM is applicable to a broad range of spatial ecological data, providing a novel, implementable tool for map comparison.

© 2016 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Ecological systems typically exhibit spatial heterogeneity arising from underlying processes that influence species occurrence, abundance, and diversity. Characterising spatial heterogeneity, and changes to it, are essential to understanding the structure of ecological systems (Fortin and Dale, 2005). Spatial ecological data range from spatially discrete events or individuals, represented as basic plots of locations in space referenced by a point (e.g. vegetation assemblages in geographical space, Penttinen et al., 1992), to distributions of species across habitats, characterised by continuous density maps (McKinney et al., 2012). Geographically referenced maps are an effective way to convey complex spatial information

because the human visual system excels at recognising structure in these familiar and intuitively read images. However, visual interpretation of spatial patterns in such maps is subjective (Da Silva-Buttkus et al., 2009), which can be further complicated by the characteristics of the mapped data, such as scale (e.g. grain and extent) and the particular cartographic representation used (e.g. projection, colour, symbology) (MacEachren, 1995). Therefore, methods have moved towards objectively quantifying the patterns observed in mapped data to produce consistent and repeatable analyses (Fortin and Dale, 2005).

The comparison of two (or more) geographically referenced maps aims to characterise differences in spatial heterogeneity and structure, and calculate defined spatial metrics between them. The problem of map comparison (Jacquez, 1995) has been studied for decades by geographers (Tobler, 1965), as well as ecologists (Levine et al., 2009). There are many ecological applications where map comparison can lead to new insights. Ecological data often have intrinsic properties that make them challenging to compare spa-

* Corresponding author at: Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews KY16 8LB, United Kingdom.

E-mail address: e1298@st-andrews.ac.uk (E.L. Jones).

tially: data tend to be continuous-valued (e.g. spatially explicit model predictions) and have underlying spatial dependencies (e.g. neighbouring cells are not independent). However, there are few established spatial comparison techniques documented in the ecological literature directly relating to the type of problems outlined above, as available methods generally only address one or other of these properties.

In recent years, emphasis has been placed on comparisons of mapped categorical data (Hagen-Zanker and Lajoie, 2008) and methods for assessing spatial structure in maps of continuous valued data or spatially explicit model predictions on a regular spatial lattice remain limited in both scope and sophistication (Hagen-Zanker, 2006a). Cell-by-cell comparisons and non-spatially explicit indexes weighted by grid cell are widely used in remote sensing, but do not account for spatial dependencies between cells (Horn, 1966; Leitão et al., 2011). Likewise, Moran's I or Geary's C tests (Cliff and Ord, 1970) assess spatial autocorrelation but provide single indices across space, which do not retain locational information. Metrics used to investigate niche similarity between species distributions predicted with Environmental Niche Models also lose spatial information to give a single measure of overlap or equivalency (Warren et al., 2008). Overlap indices and tests for spatial autocorrelation measure only one form of spatial structure in the data, and this may not be sufficient for the ecological question being posed.

A Structural Similarity Index (SSIM index) was proposed originally by Wang et al. (2004) for comparing compression techniques used in digital imaging (e.g. JPEG compression). The index uses a spatially-local moving window to generate independent components relating to local similarities in the mean, variance, and spatial correlation between the two maps being compared. SSIM can assess continuous data and simultaneously considers local magnitude and spatial structure, making it suitable to be adapted for the application of comparing spatial ecological data. Map comparison methods to ecological problems should allow uncertainty associated with the data or model predictions to be included in the map comparison to aid interpretation. Ecological maps often have uncertainty estimates associated with each grid cell when values are obtained using spatially explicit predictive models (Rocchini et al., 2011), and these should be incorporated in a map comparison approach. Additionally, local statistics such as the SSIM index are susceptible to edge effects arising from the use of a spatially local neighbourhood (Boots, 2002). Edge effects (i.e. the inclusion of null areas outside the study) are exacerbated by irregularly shaped boundaries caused by arbitrarily shaped administrative units or geographical features (e.g. islands). These may or may not influence the spatial process under study. Ecological processes often change on or near boundaries (Wiens et al., 1985), for example, the boundary of the Antarctic Circumpolar Current affects the surrounding marine ecosystem (Tynan, 1998), and so these areas can be of specific interest. Therefore, we propose two enhancements to the SSIM index to address common issues faced in spatial ecological analysis by incorporating uncertainty associated with the underlying data into the map comparison, and correcting for edge effects. We demonstrate use of the SSIM methodology and our enhancements by applying them to a case study to compare habitat preference by groups and singletons of sperm whales (*Physeter macrocephalus*, Linnaeus 1758) in the Mediterranean Sea (Pirrotta et al., 2011).

2. Methods

2.1. Map comparison

Consider two continuous valued maps (A and B) each represented as regular grids. For each cell, a local neighbourhood is defined by (n) neighbouring spatial units given a weighting (w).

The size of the neighbourhood is user-defined, has a lower limit of 3×3 cells and can take any non-even value. Wang et al. (2004) proposed the use of a (circular) Gaussian weighting function of $w = \{w_i | i = 1, 2, \dots, n\}$ where w_i is obtained from a Gaussian kernel centred on the focal cell. The standard deviation, $\sigma = n/3$, is normalised so that $\sum_{i=1}^n w_i = 1$.

The index iterates through all n cells within each local region to produce means and variances for each map as well as covariance between the two gridded maps.

$$\mu_a = \sum_{i=1}^n w_i a_i \quad (1)$$

$$\sigma_a^2 = \sum_{i=1}^n w_i (a_i - \mu_a)^2 \quad (2)$$

$$\sigma_{ab} = \sum_{i=1}^n w_i (a_i - \mu_a)(b_i - \mu_b) \quad (3)$$

μ_a , σ_a^2 , and σ_{ab} represent spatially local measures of mean, variance and covariance, computed for each cell, where a_i and b_i represent the values in cell i for maps A and B respectively. The three components of the SSIM method are then calculated from these statistics, giving spatially local measures of similarity in the mean, variance, and covariance of the two maps.

$$SIM(A, B) = \frac{2\mu_a\mu_b + c_1}{\mu_a^2 + \mu_b^2 + c_1} \quad (4)$$

$$SIV(A, B) = \frac{2\sigma_a\sigma_b + c_2}{\sigma_a^2 + \sigma_b^2 + c_2} \quad (5)$$

$$SIP(A, B) = \frac{\sigma_{ab} + c_3}{\sigma_a\sigma_b + c_3} \quad (6)$$

The statistics are named Similarity in Mean (SIM), Similarity in Variance (SIV), and Similarity in Pattern (SIP) of spatial covariance, so that they can be interpreted intuitively in ecological terms (Table 1). Constants (c_1 – c_3) are used in equations (4)–(6) to aid stability when the denominators of the equations, sum of the squared means ($\mu_a^2 + \mu_b^2$), sum of the squared variances ($\sigma_a^2 + \sigma_b^2$), and product of the standard deviations ($\sigma_a\sigma_b$) are close to 0. Following guidelines proposed by Wang et al. (2004), the constants can be estimated heuristically from the range of values of the underlying maps being compared (R) together with $k_1=0.01$ and $k_2=0.03$. Therefore, $c_1 = (k_1 R)^2$, $c_2 = (k_2 R)^2$, and $c_3 = c_2/2$.

An overall measure for comparison can be computed as the product of all three components.

$$SSIM(A, B) = [SIM(A, B)]^\alpha \cdot [SIV(A, B)]^\beta \cdot [SIP(A, B)]^\gamma \quad (7)$$

Constants α , β , γ can be used to weight individual components in SSIM and can take any value where ($\alpha > 0$, $\beta > 0$, $\gamma > 0$). Default values of $\alpha = \beta = \gamma = 1$ are used for the case study below. The overall comparison measure meets the following criteria: (1) symmetry: $SSIM(A, B) = SSIM(B, A)$; (2) boundedness: $-1 \leq SSIM(A, B) \leq 1$; and (3) unique maximum: $SSIM(A, B) = 1$ if $A = B$. SSIM is bounded by $(-1, 1)$ where -1 indicates complete dissimilarity between the spatial structure of the underlying maps, and 1 shows the maps are identical (Table 1). Similarity in pattern (SIP) of spatial covariance is of interest to ecologists because it reveals differences in spatial structure that are difficult to capture visually or through simple comparison methods such as direct map subtraction. In the context of species distribution models, low values ($SIP \rightarrow -1$) will show local-scale differences in space use that may indicate underlying mechanisms such as competitive exclusion, niche partitioning, or habitat segregation, whereas high values ($SIP \rightarrow 1$) could indicate areas where direct competition or some

Table 1

Description of local statistics calculated in SSIM index, using pairs of images (maps A and B) to demonstrate high and low similarity. The interpretation column provides a general ecological interpretation of each metric using the case example of one species in Map A and a different species in Map B.

Index	Description	Bounds	Interpretation	Map A	Map B
Similarity in means (SIM)	Ratio of twice the product of the local means to their summed squares.	(0, 1)	0 = Map A has high values; map B low values. The means are dissimilar (e.g. species have different local abundances).		
			1 = Both maps A and B have similarly high (or low) values (e.g. species have similar local abundances).		
Similarity in variance (SIV)	Ratio of twice the product of the local standard deviations to their summed variances.	(0, 1)	0 = Map A has high variance; map B low variance. The variances are dissimilar (e.g. one species is spatially clustered, the other has a homogeneous distribution).		
			1 = Both maps A and B have similarly high (or low) variance (e.g. both species have similar degrees of spatial clustering, or both have homogeneous local distributions).		
Similarity in pattern (SIP) of spatial covariance	Ratio of the local covariance to the product of the local standard deviations.	(-1, 1)	-1 = Map A has high values in some cells; Map B has high values in alternate cells. Spatial correlation is negative (e.g. species exhibit spatial partitioning).		
			0 = Map A and B exhibit no spatial correlation (e.g. species distributions are independent).		
			1 = Map A and B have high and low values in the same cells. Spatial correlation is positive (e.g. species are using the same resources, or have predator-prey interactions).		

form of ecological interaction, such as predation, are occurring. The means of each metric can be calculated to produce summary statistics (\overline{SIM} , \overline{SIV} , \overline{SIP}) if required. The mean of SSIM (\overline{SSIM}) will provide an overall metric of map comparison, capturing the similarities between means, variances, and covariance in a single value.

The mean and variance of each grid cell in the underlying maps are resampled to generate a series of realisations (N). SSIM statistics are calculated for each set of realisations ($1 \dots N$) of the two maps being compared. A variance-adjusted measure of SSIM is calculated by taking the mean of each statistic over the resulting comparisons. Upper and lower 95% confidence limits of the statistics can be calculated from the mean and variance of the sampled comparisons. To correct for edge effects, a reflection algorithm is implemented to generate synthetic buffers and ensure the spatial extent of the map comparison is preserved (Appendix A in Supplementary).

2.2. Case study: sperm whales in the Mediterranean

2.2.1. Introduction

In the Mediterranean, a small population of sperm whales persist. Sperm whales show sexually dimorphic behaviour as adults: males become increasingly solitary as they mature, and segregate from long-term social units of adult females and their offspring, excepting short term associations for mating purposes (Whitehead, 2003). It is unclear what drives this segregation and hypothe-

ses include: groups of females outcompeting solitary males when exploiting mid-water squid patches, males and females having different dietary and hence habitat preferences, or higher male growth rates that require wider search areas to locate high prey densities (Whitehead, 2003). Understanding habitat use in areas where both sexes co-occur is of obvious interest with respect to these hypotheses. Pirotta et al. (2011) predicted habitat preferences of sperm whales in the waters around the Balearic archipelago in the Mediterranean Sea. In this area both groups (assumed to be female social units) and singleton whales (assumed to be males) were regularly observed, suggesting a breeding ground. Measures of uncertainty around the estimated probability of occurrence were obtained from the modelling process to characterise variance around the mean estimate in each grid cell. Uncertainty arises in all ecological models and could result, for example, from sampling design and data collection, the resolution of environmental variables used in the model, modelling process, or the dynamic nature of species' distribution (Rocchini et al., 2011; Tessarolo et al., 2014). The SSIM index and our enhancements can be implemented accounting for any sources of uncertainty. Here, we use habitat preference maps and associated uncertainty to compare spatial patterns of use between assemblages. We demonstrate that the proposed map comparison methodology can quantify differences in the local spatial patterns observed between the maps and

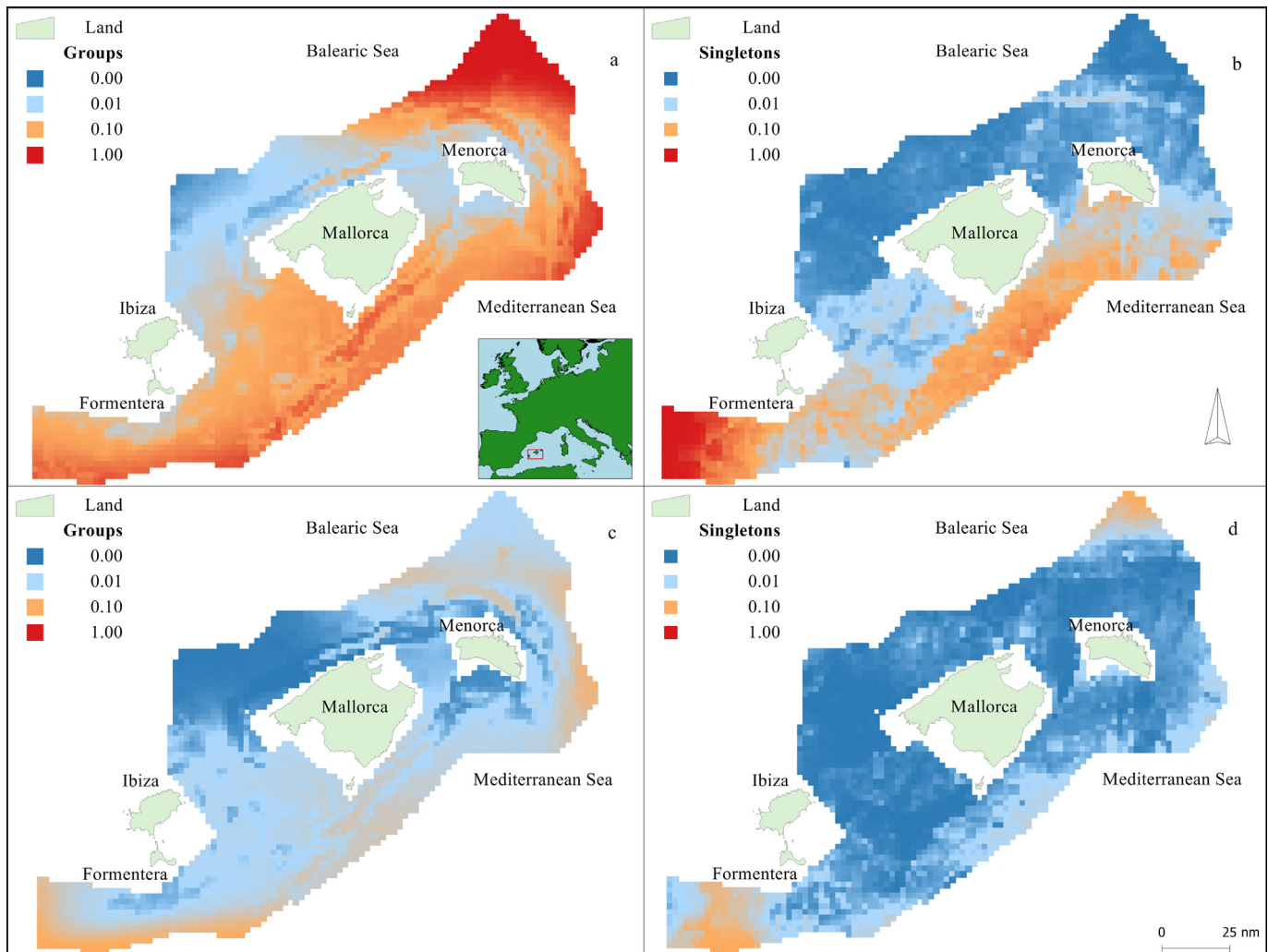


Fig. 1. Predicted probability of occurrence for both social assemblages: (a) groups; (b) singletons (c) variance in group presence; and (d) variance in singleton presence. (a) and (b) are modified from [Pirrotta et al. \(2011\)](#). Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG) shoreline data from NOAA were used, available to download from <http://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html> <http://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html>.

provide novel biological insights not readily apparent from visual assessments alone.

2.2.2. Data

The study area was located from 38 to 41°N and 0.5–5°E, centred on the islands of Ibiza, Mallorca and Menorca ([Fig. 1](#)). Information on sperm whale occurrence was collected during dedicated summer research cruises covering the waters around the Balearic archipelago. Each cruise lasted for approximately a month and was repeated over 6 consecutive years (2003–2008). Whales were located and tracked acoustically from their echolocation clicks, and an encounter was defined as a period of continuous acoustic contact with one or more animals. [Pirrotta et al. \(2011\)](#) used a Generalised Additive Modelling (GAM) approach to model sperm whale occurrence as a function of several environmental and temporal predictors, combined with Generalised Estimating Equations (GEEs) to account for autocorrelation in the residuals. Further details on the environmental datasets and analytical approach can be found in [Pirrotta et al. \(2011\)](#) and are summarised here: Separate analyses were carried out for singletons and groups to determine whether habitat preference was characterised by different extrinsic drivers. The final model for sperm whale groups included latitude, longitude, weekly sea surface temperature (SST) and slope gradient. For singletons, latitude, longitude, year, monthly SST and slope aspect were retained by model selection. The authors noted

qualitatively different spatial patterns emerging for the two social assemblages in the final prediction maps, quantitatively supported by an inverse relationship with SST, and suggested that these might be the result of fine-scale habitat segregation.

2.2.3. Analysis

The predicted probability of presence of groups and singletons, and corresponding estimates of variance were mapped at a spatial resolution of 2 nautical miles (NM) on a regular grid. To calculate SSIM statistics, the size of the local neighbourhood for both maps should be defined by taking the nature of the underlying data and ecological process in question into account. [Lewis et al. \(2007\)](#) examined the nearest-neighbour distances between sperm whales in the Mediterranean Sea using a similar acoustic survey approach to data used in [Pirrotta et al. \(2011\)](#) and found that animals defined as belonging to a ‘cluster’ mostly had an upper limit of 2.7 NM of perpendicular distance between them, whereas dispersed (singleton) animals were separated by distances beyond this threshold. The size of the local neighbourhood was defined in a 3×3 ($n=9$) cell window (6 NM x 6 NM), such that the edge of the window was at least 2 NM (1 grid cell) from any animals encountered in the centre cell. A circular Gaussian weighting kernel ($w = \{w_i | i = 1, 2, \dots, 9\}$) with a standard deviation ($\sigma = n/3$) was set to 3 NM. Sensitivity tests were applied to verify that varying the size of local neighbourhood and using a Gaussian weighting kernel did not affect results

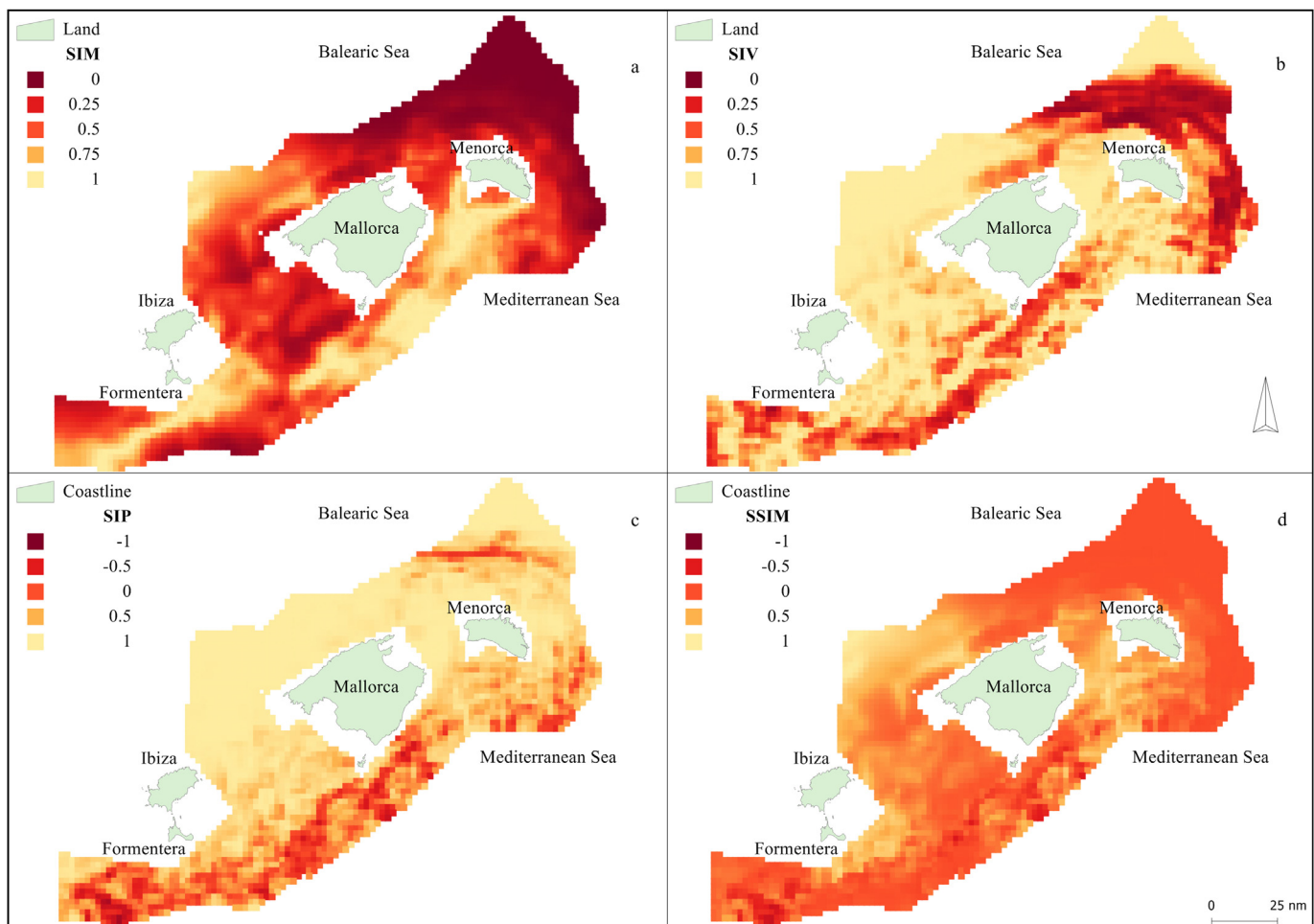


Fig. 2. Map comparison between the predicted probability of occurrences of group and singleton sperm whales. A reflection algorithm was used to counteract internal edge effects: (a) Similarity In Means (0–1); (b) Similarity In Variance (0–1); (c) Similarity In Pattern of spatial covariance (–1 to 1); and (d) Structural Similarity index (–1 to 1).

from the case study (Section 2.2.4 and Appendix B in Supplementary). Uncertainty from the underlying data was included in the map comparisons using parametric bootstrapping. Samples from a multivariate normal distribution were generated using model coefficients and each covariance matrix to produce 500 realisations of model coefficients for the group and singleton models. These were used to predict 500 sets of probabilities for the group and singleton models. SSIM statistics were calculated for each pair of maps generated from bootstrapped data. Mean and variance of predicted probabilities in each grid cell were taken for each statistic (SIM, SIV, SIP, and SSIM). All analysis was conducted using the statistical software package R (R Core Team, 2014), and code and data used for calculating the SSIM index can be obtained from Appendices C and D in Supplementary respectively.

2.2.4. Sensitivity testing

Sets of sensitivity tests were conducted to demonstrate how varying specific (user-defined) parameters could potentially affect results of the map comparison analysis for the sperm whale data: (1) A circular Gaussian weighting kernel was applied to the local neighbourhood window vs. no weighting; (2) the size of the local neighbourhood was varied, using 3×3 , 5×5 , and 7×7 grid cells; and (3) a reflection algorithm to correct for edge effects was applied vs. no edge correction. For Gaussian weighting tests, the size of the local neighbourhood (w) was set at 3×3 grid cells and the reflection algorithm was implemented. For local neighbourhood tests, Gaussian weighting and the reflection algorithm were applied. For edge

effects tests, the size of the local neighbourhood (w) was set at 3×3 grid cells and Gaussian weighting was applied. In all tests, $\sigma = w/3$ and only mean values from the underlying maps being compared were used. SSIM statistics were calculated for each set of tests and means and variances of each statistic (SIM, SIV, SIP, SSIM) were calculated to provide summary statistics. Welch two-sample t -tests were used to compare the SSIM statistic for each set of tests.

3. Results

SSIM was used to compare the predicted probability of occurrence between groups and singleton sperm whales (Fig. 2). Fig. 2a, showing similarity between the local means (SIM), aligns with visual differences seen between the underlying maps (Figs. 1a and b). Areas where SIM is close to 1 (yellow) are found in regions of the study area where habitat preference is high for both social assemblages (east and south of Mallorca, and east and south of Formentera), or low for both social assemblages (north-west of Mallorca). Values of SIM close to 0 (red) denote areas where one social assemblage has low habitat preference and the other has high habitat preference. An example is to the north of Menorca, where fewer data were collected: the sperm whale group model in particular was subject to sampling bias, resulting in a high estimate for habitat preference, whereas predicted probability of occurrence for singletons was low (Pirodda et al., 2011). Fig. 2b shows similarity in local variance (SIV) between the maps. Values close to 1 (yellow) show areas where assemblages have similar variance in the

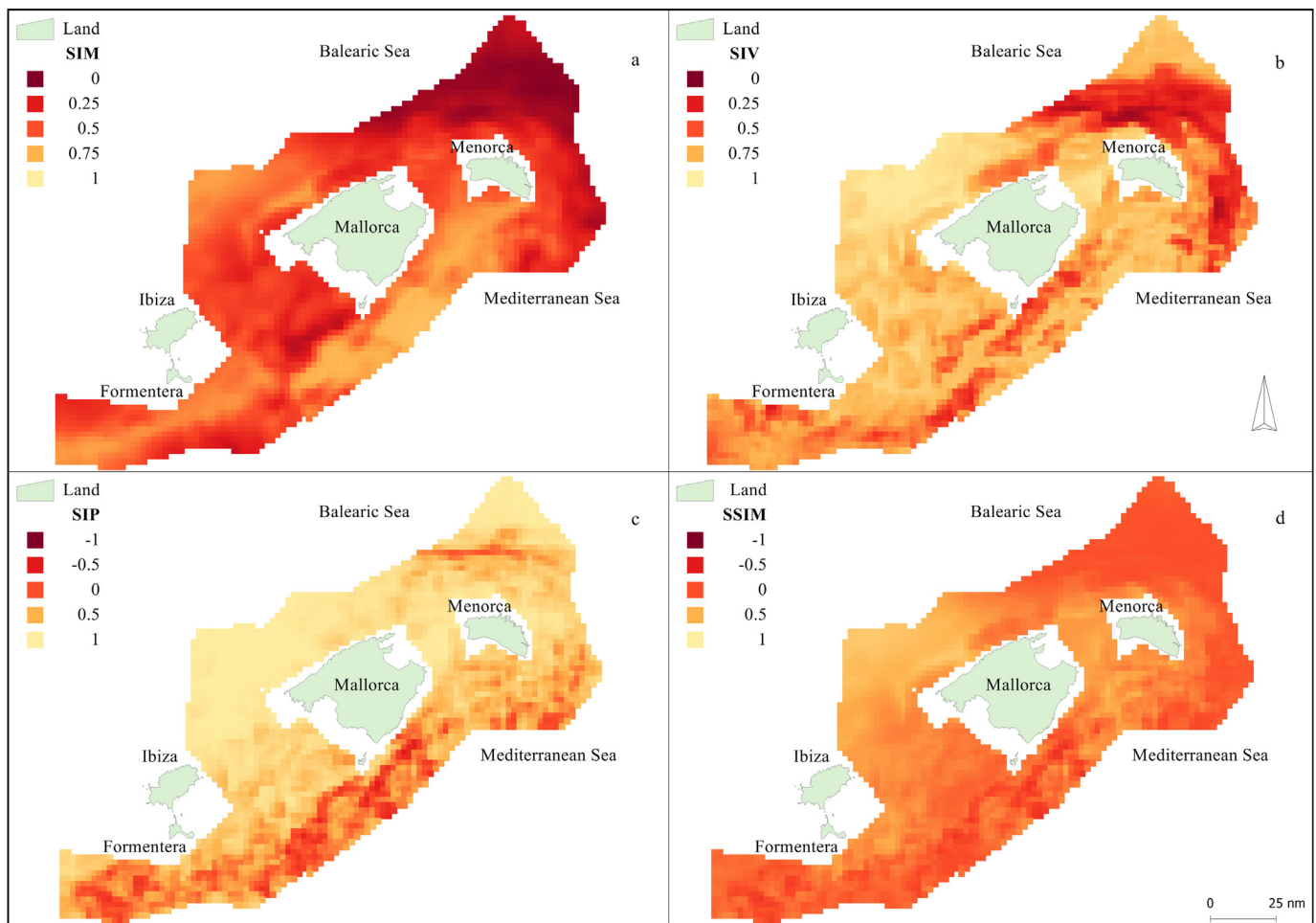


Fig. 3. Map comparison between the predicted probability of occurrence of group and singleton sperm whales using bootstrapped uncertainty from the habitat preference models of both social assemblages. A reflection algorithm was used to counteract internal edge effects: (a) Similarity In Means (0–1); (b) Similarity In Variance (0–1); (c) Similarity In Pattern of spatial covariance (–1 to 1); and (d) Structural Similarity index (–1 to 1).

probability of occurrence, and values close to 0 (red) show areas where the variance is different. For example, the area to the north of Menorca shows a transition zone where groups and singletons are using space differently—groups have heterogeneous, sporadic space use (i.e. high variance), singletons are utilising space in a consistent, homogeneous way (i.e. low variance). Fig. 2c shows the similarity in patterns (SIP) of spatial covariance between the maps. The SIP metric is the most difficult to capture through visual comparison of habitat use between groups and singletons (c.f. Figs. 1a and b). Values close to 1 (yellow) denote local regions where the spatial structure between predicted probability of occurrence of groups and singletons is similar, meaning grid cells with relative high and low variance are in the same locations in each underlying map. Underlying mechanisms of direct competition for resources could be occurring, for example to the north, east and west of Menorca, and north and west of Mallorca. Values close to –1 (red) indicate areas where local spatial structure is dissimilar, suggesting spatial partitioning may be occurring (north of Menorca in the transition zone discussed previously, and the southern edge of the study area). Fig. 2d shows SSIM, which is the product of the other three statistics. Differences in spatial structure detected in SIV (Fig. 2b) and SIP (Fig. 2c) at the southern edge of the study area remain apparent in the SSIM index. Some spatial structural similarities seen throughout Figs. 2a–c to the north-west and east of Mallorca, and south-east of Formentera are also retained in SSIM. The mean

value of SSIM was calculated ($\overline{SSIM} = 0.22$), showing positive spatial structure between the underlying maps.

Results incorporating uncertainty from the underlying maps into the comparison calculation are provided in Fig. 3, and show similar inferences to those in Fig. 2, although each of the four comparison metrics exhibit less extreme values. An area of particular interest is south of Mallorca (Fig. 3c), where SIP is close to –1 (red), characterising different spatial patterns in habitat use between groups and singletons. Fig. 4 focuses on this area, which is situated over the continental slope and has previously been identified as a feeding ground for sperm whales (Gannier and Praca, 2007; Gannier et al., 2002). Although Fig. 4a shows that both social assemblages have similar (high) habitat preference (SIM is close to 1), there is strong negative SIP in specific areas (Fig. 4c), indicating local-scale spatial partitioning between groups and singletons. These patterns occur mostly along bathymetric contours at depths ranging between 1000 to 2000 m. The spatial structure can be seen in SSIM (Fig. 4d).

The results of sensitivity tests are shown in Table 2 (visual results are available in Appendix B in Supplementary). When comparing Gaussian weighting vs. no weighting, there was no significant difference in \overline{SSIM} ($t = 0.06$, p -value = 0.95). By varying the size of the local neighbourhood between 3×3 and 5×5 grid cells SIV and SIP show differences (calculated from the variance and covariance in the underlying maps being compared), leading to a significant t-test result when comparing \overline{SSIM} ($t = 4.14$, p -value < 0.00005). Likewise,

Table 2
Results determining sensitivity of the SSIM index when user-defined parameters were varied: applying Gaussian weighting to the local neighbourhood vs. no weighting; varying the size of the local neighbourhood; and implementing the reflection algorithm vs. not centering for edge effects.

Sensitivity analysis	Parameters	SIV			SIP			SSIM			t-test on SSIM	
		Mean	Low 95%CI	Upp 95%CI	Mean	Low 95%CI	Upp 95%CI	Mean	Low 95%CI	Upp 95%CI	t	p-value
Weighting	Gaussian weighting	0.432	-0.183	1.048	0.741	0.169	1.313	0.654	-0.127	1.434	0.218	-0.293
	No Gaussian weighting	0.432	-0.183	1.048	0.740	0.167	1.313	0.652	-0.130	1.435	0.218	-0.294
Local neighbourhood	3 × 3 cells	0.432	-0.183	1.048	0.741	0.169	1.313	0.654	-0.127	1.434	0.218	-0.293
	5 × 5 cells	0.433	-0.174	1.041	0.706	0.104	1.308	0.567	-0.229	1.362	0.191	-0.284
	7 × 7 cells	0.435	-0.169	1.039	0.684	0.063	1.306	0.515	-0.284	1.314	0.177	-0.279
Edge effects	Reflection algorithm	0.432	-0.183	1.048	0.741	0.169	1.313	0.654	-0.127	1.434	0.218	-0.293
	No reflection algorithm	0.435	-0.177	1.047	0.734	0.155	1.314	0.639	-0.150	1.427	0.212	-0.287

^a 5 × 5 Window vs. 7 × 7 were compared to give this t-test result.

when comparing 5 × 5 and 7 × 7 tests, there is a significant difference between $SSIM(t = 2.34, p\text{-value} = 0.02)$. When the reflection algorithm was not applied, the value of $SSIM$ was not affected significantly but there was a reduction in the spatial extent of the map comparison (as values for edge cells could not be calculated) (Appendix B in Supplementary).

4. Discussion

We have described an approach to objectively compare spatial patterns between two continuous valued maps. We enhanced the original SSIM index (Wang et al., 2004) by incorporating uncertainty from underlying maps into the comparison calculation and correcting for edge effects. Application of the SSIM approach, including our enhancements, was demonstrated with a case study using sperm whale distribution data in the Mediterranean Sea. Quantitative map comparison tools are currently limited in their extent and application in the ecological literature (Hagen-Zanker, 2006b; Robertson et al., 2014), possibly because ecological data have characteristic properties such as continuous values and inherent spatial dependencies that make quantifying the underlying spatial structure between geographically referenced maps challenging. As well as accounting for these characteristics, the SSIM index has several key advantages making it ideal for broader ecological applications. First, the methodology can be easily implemented regardless of the prediction or estimation method used to obtain the underlying maps. For instance, a useful application of the method would be to compare two maps where different statistical methods were used to address similar questions. Second, the SSIM index produces a number of underlying statistics, as well as an overall measure of similarity in spatial structure. By comparing local means, variances, and covariance between underlying maps, different aspects of spatial patterns are characterised, potentially providing insight into underlying processes that drive these patterns. Finally, the size of the local neighbourhood in the map comparison calculation is user-defined. Prior knowledge of spatial scale of the data can be used to inform the map comparison analysis, providing more meaningful results.

Dependent on the size of the local neighbourhood, edge effects occur when comparing maps because non-valued cells beyond the boundary of the study area are included. To ensure the map comparison produced the same spatial extent as the underlying maps, a reflection algorithm was chosen to correct for edge effects because of its ability to deal with complex edges and ease of implementation. The algorithm reflected known data along edges to extrapolate outside of the study area. A limitation of this method is that it can emphasise fine-scale or local patterns in areas where it is implemented, and so care should be taken when interpreting results close to edges in the study area.

The definition of spatially local neighbourhoods and the effects of their size have been well studied (Chefaoui, 2014; Long et al., 2010; Zurlini et al., 2007). In ecology, local neighbourhood size must be considered in the context of spatial resolution of the data and the underlying ecological processes being investigated (Wiens, 1989; Wu, 2004). Therefore, local neighbourhood sizes are often varied to examine their influence on results and inferences. However, there can still be subjectivity in selecting the appropriate local scale for spatial analysis (Nelson and Boots, 2008). In the case study, a local neighbourhood was selected based on the spatial scale present in the ecological process (i.e. the distance between the assemblages being compared). As the local neighbourhood size increases fine-scale differences in patterns identified through SSIM will disappear, resulting in a smoothing effect. Similarly, using the minimum local neighbourhood (3 × 3 grid squares) may produce results that show fine-scale differences in patterns which do not make sense ecolog-

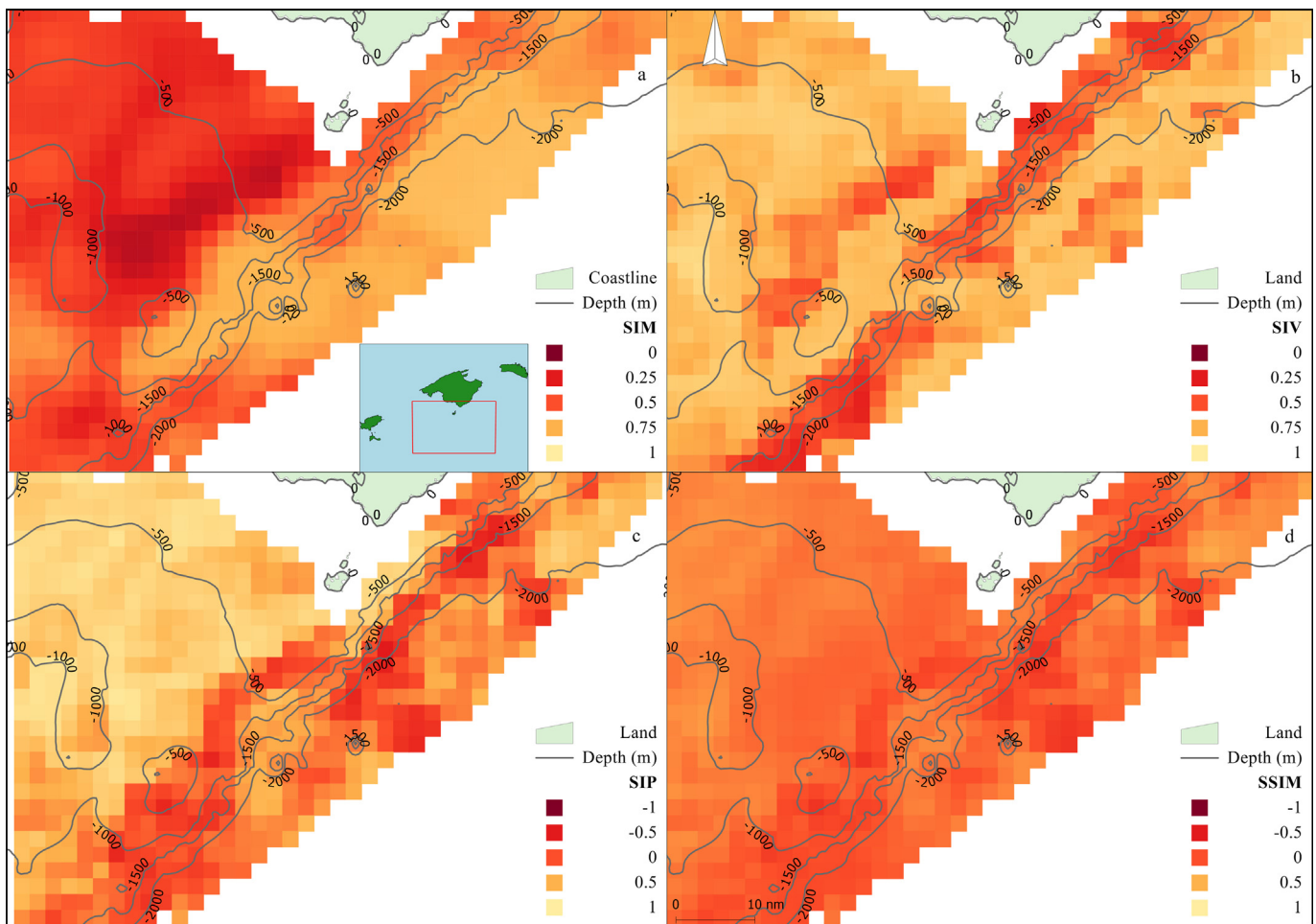


Fig. 4. Map comparison between the predicted probability of occurrence of group and singleton sperm whales using bootstrapped uncertainty from the habitat preference models of both social assemblages. A reflection algorithm was used to counteract internal edge effects. Focusing on the area of interest south of Mallorca: (a) Similarity In Means (0–1); (b) Similarity In Variance (0–1); (c) Similarity In Pattern of spatial covariance (–1 to 1); and (d) Structural Similarity index (–1 to 1).

ically. We recommend use of a local neighbourhood in the SSIM index that bears relation to: (1) the scale of underlying data (e.g. movement of animals) represented in the maps being compared, (2) ecological questions being investigated through map comparisons, and, (3) scale-dependent patterns of underlying ecological processes being investigated.

4.1. Case study

The map comparison showed differences in space use between groups and singleton sperm whales. The area south of Mallorca was also identified by [Pirodda et al. \(2011\)](#) as important to both social assemblages. Data sampling effort was greatest here and therefore groups and singleton models suffered less from sampling bias than in other areas. In this area, both social assemblages had high predicted probability of occurrence. Statistical differences in patterns of space-use were associated with fine-scale features at depths between 1000 and 2000 m, and the probability of occurrence for both social assemblages was previously found to be driven by bathymetric features ([Pirodda et al., 2011](#)). Although the sex of animals included in the study was not verified, singletons showed diving behaviour typical of solitary males, and groups of sperm whales are generally associated with adult females and their immature offspring ([Drouot et al., 2004](#); [Whitehead, 2003](#)). [Whitehead \(2003\)](#) suggested that reduced foraging success for males in areas where both social assemblages exist may be a result of resource

competition. Our results reveal that groups and singletons do interact spatially in some mutually exclusive way. This has implications for both understanding local space use, and informing more general hypotheses about the evolution of extreme behavioural sexual dimorphism in sperm whales ([Whitehead and Weilgart, 2000](#)). The results provide a specific target area so that efficient resources can be put into studying sexual segregation of groups and singletons. Hypotheses could be investigated to determine whether patterns of mutually exclusive (presumed) foraging of groups and individuals show stable resource partitioning (in which case both social assemblages may be foraging optimally), or whether patterns are the result of one social assemblage being outcompeted and forced to utilise sub-optimal habitat. Sensitivity tests indicated that comparison results were affected by neighbourhood size, and any interpretation should take account of this.

4.2. Broader applications and further development

Comparisons of spatially referenced data provide a mechanism for linking observed spatial patterns with underlying ecological processes ([Turner, 1989](#)). Methodology presented here has wider applications for ecology, where quantitative comparisons of spatial patterns are often required to understand underlying processes and guide management decisions. Application of the SSIM index with our enhancements provides spatially explicit comparisons to identify areas where there are underlying differences in space-

use. There are many applications to spatial ecology problems such as identifying areas of conflict between anthropogenic activities and wildlife: depredation on domestic livestock and farmed species by apex predators (Berland et al., 2008; Ripple et al., 2014; Suryawanshi et al., 2013). An important application is the assessment of change in distribution between and within species, such as comparing density maps obtained using different methods (Bailey et al., 2014), assessing competition and spatial segregation between species (Suryawanshi et al., 2013; Wilson, 2010), and seasonal changes in distribution (Millspaugh et al., 2015).

The Marine Strategy Framework Directive uses an ecosystem-based approach to management of anthropogenic activities within the marine environment (Olenin et al., 2010). Under this framework, ecosystems are assessed through a set of environmental abundance and distribution indicators to determine conservation status. Spatially-explicit indicators such as biodiversity indices (species richness and diversity) present mean values over time (Piroddi et al., 2015). The SSIM index and enhancements presented here can be used to elevate these indicators to a spatio-temporal context and assess biodiversity over time. When used in conjunction with abundance estimates, these can further inform the spatial management process.

The methodology could be developed further. Currently, spatial resolution and extent of the maps being compared must be regularly spaced and identical, and the case study used to demonstrate the methodology benefitted from having these characteristics. However, comparing animal distributions (e.g. using line transect data from animal sightings surveys) may result in varying spatial resolution and extent because sampling effort and survey area can change over time. Adapting SSIM methodology to allow for maps with non-regular lattices and point-process patterns to be compared, would be beneficial for effective analyses. For longer time-series (Bailey et al., 2014) or multiple species comparisons (Wilson, 2010), map comparison functionality could be extended to compare more than two maps at once, either sequentially, or through pair-wise comparisons.

5. Conclusions

The SSIM index and enhancements presented here offer a comprehensive tool to objectively compare spatially explicit ecological data within an implementable framework. An advantage of the SSIM index is that different aspects of spatial comparison can be investigated: maps of SIM, SIV, and SIP (relating to similarities in local means, variances, and covariance, respectively) can be calculated to reveal spatial patterns that cannot be seen through visual inspection of the underlying maps. The SSIM metric summarises SIM, SIV, and SIP into one map because summary statistics are often required to condense information. This can be further summarised by calculating the mean over SSIM to give a single value representing similarity between the underlying maps.

We presented enhancements to the SSIM index by incorporating uncertainty from the underlying maps and correcting for edge effects so that the methodology can be broadly applied to many types of spatial ecological data. Using an ecological case study to compare groups and singletons sperm whale distribution in the Mediterranean Sea, we demonstrated the presence of local-scale spatial structure that could not be detected either visually or using map subtraction techniques. We found that in these areas where (presumed) foraging was taking place, singletons and groups of whales were spatially mutually exclusive. This enabled us to recommend that future behavioural studies focusing on interactions between singletons and groups of whales whilst foraging could most effectively be carried out in the areas of interest we have identified.

Acknowledgements

E.L.J. was funded under Scottish Government grant MMSS001/11. Sperm whale data were collected with support from One World Wildlife, the Natural Environment Research Council (NER/I/S/2002/00632), Whale and Dolphin Conservation (WDC), and J.M. Brotons of the Balearic Government Office of Fisheries Management. L.R. was supported by the MASTS pooling initiative, funded by the Scottish Funding Council (HR09011) and contributing institutions and their support are gratefully acknowledged. Thank you to J.B. Illian, G.D. Ruxton, and two anonymous reviewers for reviewing an earlier version of the manuscript and providing valuable feedback. This manuscript partially fulfils the PhD submission of E.L.J., part-funded by CREEM, University of St Andrews.

Appendices. Supplementary Material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ecolind.2016.05.051>.

References

- Bailey, H., Hammond, P.S., Thompson, P.M., 2014. Modelling harbour seal habitat by combining data from multiple tracking systems. *J. Exp. Mar. Biol. Ecol.* 450, 30–39, <http://dx.doi.org/10.1016/j.jembe.2013.10.011>.
- Berland, A., Nelson, T., Stenhouse, G., Graham, K., Cranston, J., 2008. The impact of landscape disturbance on grizzly bear habitat use in the Foothills Model Forest, Alberta, Canada. *For. Ecol. Manage.* 256, 1875–1883, <http://dx.doi.org/10.1016/j.foreco.2008.07.019>.
- Boots, B., 2002. Local measures of spatial association. *Ecoscience* 9, 168–176.
- Chefaoui, R.M., 2014. Landscape metrics as indicators of coastal morphology: a multi-scale approach. *Ecol. Indic.* 45, 139–147, <http://dx.doi.org/10.1016/j.ecolind.2014.04.004>.
- Cliff, A.D., Ord, K., 1970. Spatial autocorrelation: a review of existing and new measures with applications. *Econ. Geogr.* 46, 269–292.
- Da Silva-Buttkus, P., Marcelli, G., Franks, S., Stark, J., Hardy, K., 2009. Inferring biological mechanisms from spatial analysis: prediction of a local inhibitor in the ovary. *Proc. Natl. Acad. Sci. U. S. A.* 106, 456–461, <http://dx.doi.org/10.1073/pnas.0810012106>.
- Drouot, V., Gannier, A., Goold, J.C., 2004. Summer social distribution of sperm whales (*Physeter macrocephalus*) in the Mediterranean Sea. *J. Mar. Biol. Assoc. UK* 84, 675–680, <http://dx.doi.org/10.1017/S0025315404009749h>.
- Fortin, M.J., Dale, M.R.T., 2005. *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press, Cambridge.
- Gannier, A., Praca, E., 2007. SST fronts and the summer sperm whale distribution in the north-west Mediterranean Sea. *J. Mar. Biol. Assoc. UK* 87, 187–193, <http://dx.doi.org/10.1017/S0025315407054689>.
- Gannier, A., Drouot, V., Goold, J.C., 2002. Distribution and relative abundance of sperm whales in the Mediterranean Sea. *Mar. Ecol. Prog. Ser.* 243, 281–293.
- Hagen-Zanker, A., Lajoie, G., 2008. Neutral models of landscape change as benchmarks in the assessment of model performance. *Landscape Urban Plann.* 86, 284–296, <http://dx.doi.org/10.1016/j.landurbplan.2008.04.002>.
- Hagen-Zanker, A., 2006a. Comparing continuous Valued Raster Data: A Cross Disciplinary Literature Scan, Netherlands Environmental Assessment Agency. Research Institute for Knowledge systems, Maastricht, The Netherlands.
- Hagen-Zanker, A., 2006b. Map comparison methods that simultaneously address overlap and structure. *J. Geogr. Syst.* 8, 165–185, <http://dx.doi.org/10.1007/s10109-006-0024-y>.
- Horn, H.S., 1966. Measurement of overlap in comparative ecological studies. *Am. Nat.* 100, 419–424, <http://dx.doi.org/10.2307/2678832>.
- Jacquez, G.M., 1995. The map comparison problem: tests for the overlap of geographic boundaries. *Stat. Med.* 14, 2343–2361.
- Leitão, P.J., Moreira, F., Osborne, P.E., 2011. Effects of geographical data sampling bias on habitat models of species distributions: a case study with steppe birds in southern Portugal. *Int. J. Geogr. Inf. Sci.* 25, 439–454, <http://dx.doi.org/10.1080/13658816.2010.531020>.
- Levine, R.S., Yorita, K.L., Walsh, M.C., Reynolds, M.G., 2009. A method for statistically comparing spatial distribution maps. *Int. J. Health Geogr.* 8, 7, <http://dx.doi.org/10.1186/1476-072X-8-7>.
- Lewis, T., Gillespie, D., Lacey, C., Matthews, J., Danbolt, M., Leaper, R., McLanaghan, R., Moscrop, a., 2007. Sperm whale abundance estimates from acoustic surveys of the Ionian Sea and Straits of Sicily in 2003. *J. Mar. Biol. Assoc. UK* 87, 353–357, <http://dx.doi.org/10.1017/S0025315407054896>.
- Long, J.A., Nelson, T.A., Wulder, M.A., 2010. Local indicators for categorical data: impacts of scaling decisions. *Can. Geogr.* 54, 15–28, <http://dx.doi.org/10.1111/j.1541-0064.2009.00300.x>.

- MacEachren, A.M., 1995. *How Maps Work: Representation, Visualization, and Design*. The Guilford Press, New York.
- McKinney, J., Hoffmayer, E., Wu, W., Fulford, R., Hendon, J., 2012. Feeding habitat of the whale shark *Rhincodon typus* in the northern Gulf of Mexico determined using species distribution modelling. *Mar. Ecol. Prog. Ser.* 458, 199–211, <http://dx.doi.org/10.3354/meps09777>.
- Millspaugh, J.J., Rittenhouse, C.D., Montgomery, R. a., Matthews, W.S., Slotow, R., 2015. Resource selection modeling reveals potential conflicts involving reintroduced lions in Tembe Elephant Park, South Africa. *J. Zool.* 296, 124–132, <http://dx.doi.org/10.1111/jzo.12224>.
- Nelson, T., Boots, B., 2008. Detecting spatial hot spots in landscape ecology. *Ecography* 31, 556–566.
- Olenin, S., Alemany, F., Cardoso, A., Gollasch, S., Goulletquer, P., Lehtiniemi, M., McCollin, T., Minchin, D., Miossec, L., Ambrogì, A., Ojaveer, H., 2010. Marine Strategy Framework Directive. Task Gr. 2.
- Penttinen, A., Stoyan, D., Henttonen, H.M., 1992. Marked point processes in forest statistics. *For. Sci.* 38, 806–824.
- Piroddi, C., Teixeira, H., Lynam, C.P., Smith, C., Alvarez, M.C., Mazik, K., Andonegi, E., Churilova, T., Tedesco, L., Chifflet, M., Chust, G., Galparsoro, I., Garcia, A.C., Kämäri, M., Kryvenko, O., Lassalle, G., Neville, S., Niquil, N., Papadopoulou, N., Rossberg, A.G., Suslin, V., Uyarra, M.C., 2015. Using ecological models to assess ecosystem status in support of the European Marine Strategy Framework Directive. *Ecol. Indic.* 58, 175–191, <http://dx.doi.org/10.1016/j.ecolind.2015.05.037>.
- Pirotta, E., Matthiopoulos, J., Mackenzie, M.L., Scott-Hayward, L., Rendell, L., 2011. Modelling sperm whale habitat preference: a novel approach combining transect and follow data. *Mar. Ecol. Prog. Ser.* 436, 257–272, <http://dx.doi.org/10.3354/meps09236>.
- R Core Team, 2014. R: A language and environment for statistical computing.
- Ripple, W.J., Estes, J.A., Beschta, R.L., Wilmers, C.C., Ritchie, E.G., Hebblewhite, M., Berger, J., Elmhagen, B., Letnic, M., Nelson, M.P., Schmitz, O.J., Smith, D.W., Wallach, A.D., Wirsing, A.J., 2014. Status and ecological effects of the world's largest carnivores. *Science* 343, <http://dx.doi.org/10.1126/science.1241484> (1241484).
- Robertson, C., Long, J.A., Nathoo, F.S., Nelson, T.A., Plouffe, C.C.F., 2014. Assessing quality of spatial models using the structural similarity index and posterior predictive checks. *Geogr. Anal.* 46, 53–74, <http://dx.doi.org/10.1111/gean.12028>.
- Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jimenez-Valverde, a., Ricotta, C., Bacaro, G., Chiarucci, a., 2011. Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Prog. Phys. Geogr.* 35, 211–226, <http://dx.doi.org/10.1177/0309133311399491>.
- Suryawanshi, K.R., Bhatnagar, Y.V., Redpath, S., Mishra, C., 2013. People, predators and perceptions: patterns of livestock depredation by snow leopards and wolves. *J. Appl. Ecol.* 50, 550–560, <http://dx.doi.org/10.1111/1365-2664.12061>.
- Tessarolo, G., Rangel, T.F., Araujo, M.B., Hortal, J., 2014. Uncertainty associated with survey design in species distribution models. *Divers. Distrib.* 20, 1258–1269, <http://dx.doi.org/10.1111/ddi.12236>.
- Tobler, W.R., 1965. Computation of the correspondence of geographical patterns. *Pap. Reg. Sci. Assoc.* 15, 131–139.
- Turner, M.G., 1989. Landscape ecology: the effect of pattern on process. *Annu. Rev. Ecol. Syst.* 20, 171–197, <http://dx.doi.org/10.1146/annurev.es.20.110189.001131>.
- Tynan, C.T., 1998. Ecological importance of the southern boundary of the Antarctic circumpolar front. *Nat. Lett.* 392, 708–710, <http://dx.doi.org/10.1016/j.infsof.2008.09.005>.
- Wang, Z., Bovik, A.C., Sheikh, H.R., Member, S., Simoncelli, E.P., Member, S., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13, 600–612.
- Warren, D.L., Glor, R.E., Turelli, M., 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* (N. Y.) 62, 2868–2883, <http://dx.doi.org/10.1111/j.1558-5646.2008.00482.x>.
- Whitehead, H., Weilgart, L., 2000. *Cetacean societies: field studies of dolphins and whales*. In: Mann, J., Connor, R., Tyack, P., Whitehead, H. (Eds.), *Cetacean Societies: Field Studies of Dolphins and Whales*. University of Chicago Press, Chicago, IL, USA, pp. 154–172.
- Whitehead, H., 2003. *Sperm Whales: Social Evolution in the Ocean*. University of Chicago Press, Chicago, IL, USA.
- Wiens, J.A., Crawford, C.S., Gosz, J.R., 1985. Boundary dynamics: a conceptual framework for studying landscape ecosystems. *Oikos* 45, 421–427, <http://dx.doi.org/10.2307/3565577>.
- Wiens, J.A., 1989. Spatial scaling in ecology. *Funct. Ecol.* 3, 385–397, <http://dx.doi.org/10.2307/2389612>.
- Wilson, R.P., 2010. Resource partitioning and niche hyper-volume overlap in free-living Pygoscelid penguins. *Funct. Ecol.* 24, 646–657, <http://dx.doi.org/10.1111/j.1365-2435.2009.01654.x19>.
- Wu, J., 2004. Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecol.* 19, 125–138, <http://dx.doi.org/10.1023/B:LAND.0000021711.40074.ae>.
- Zurlini, G., Riitters, K.H., Zaccarelli, N., Petrosillo, I., 2007. Patterns of disturbance at multiple scales in real and simulated landscapes. *Landscape Ecol.* 22, 705–721, <http://dx.doi.org/10.1007/s10980-006-9055-5>.